DOCUMENT RESUME

ED 232 567 HE 016 491

AUTHOR Lay, Robert S.
TITLE New Software for Market Segmentation Analysis: A Chi-Square Interaction Detector. AIR 1983 Annual Forum Paper.
PUB DATE May 83
NOTE 27p.; Paper presented at the Annual Forum of the Association for Institutional Research (Toronto, Ontario, May 23-26, 1983). Research supported in part by a grant from the National Association of College Admissions Counselors.
PUB TYPE Viewpoints (120) -- Reports - Descriptive (141)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *College Applicants; *Computer Programs; Evaluation Methods; Flow Charts; Higher Education; *Institutional Research; Marketing; Nonparametric Statistics; *Prediction; *Statistical Analysis
IDENTIFIERS *AIR Forum; *Chi Square Analysis

ABSTRACT

The advantages and disadvantages of new software for market segmentation analysis are discussed, and the application of this new, chi-square based procedure (CHAID), is illustrated. A comparison is presented of an earlier, binary segmentation technique (THAID) and a multiple discriminant analysis. It is suggested that CHAID is superior to earlier data analysis techniques such as the Automatic Interaction Detector (AID) in some important respects. CHAID uses a chi-square statistic to measure the strength of association between a given predictor and the criterion to be maximized. Because the chi-square statistic is nonparametric, CHAID requires fewer assumptions than does AID. Many of the advantages of CHAID stem from its use of a chi-square optimizing procedure, which is useful for segmenting applicant and prospective applicant pools. For college planning, the approach serves as a guide for distributing scarce resources. The example of a segmentation analysis identifies prospective students who have some chance of being admitted, based on questionnaire data. The CHAID technique is shown to have additional flexibility, both in the variety of variables that may be optimized and in the way the technique can identify multiple segments at any given step. (SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

#10

ED232567

New Software for Market Segmentation Analysis: A Chi-Square Interaction Detector

Robert S. Lay

Director
Enrollment Management Research
Boston College
Lyons Hall, 106
Chestnut Hill, MA 02167
(617)969-0100, x3297

Presented to
the Twenty-Third Annual Forum of
The Association for Institutional Research
Toronto, Ontario
May 23-26, 1983

Research supported in part by a grant from the National Association of College Admissions Counselors.

HE016491

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

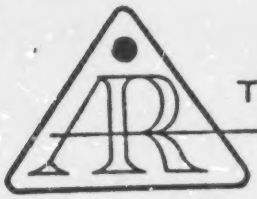
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

AIR

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."



THE ASSOCIATION FOR INSTITUTIONAL RESEARCH

This paper was presented at the Twenty-Third Annual Forum of the Association for Institutional Research held at the Sheraton Centre in Toronto, Ontario, May 23-26, 1983. This paper was reviewed by the AIR Forum Publications Committee and was judged to be of high quality and of interest to others concerned with the research of higher education. It has therefore been selected to be included in the ERIC Collection of Forum papers.

D. R. Coleman, Chairman
Forum Publication
Advisory Committee

ABSTRACT

New Software for Market Segmentation Analysis:
A Chi-Square Interaction Detector

Over the last twenty years, a revolution in statistical methods has transformed marketing research. As with any revolution, it has taken some time for the new ways to become recognized as useful and for the supporting technologies to be elaborated and adopted by practitioners. In this article, we discuss the advantages and disadvantages of newly developed, readily available software that improves upon a popular method of exploratory data analysis first introduced in the early 1960's. We illustrate how this new, chi-square based procedure (CHAID) optimally segments a population along the dimension of a desired outcome--application to a university in our example. An earlier, binary segmentation technique, AID and a multiple discriminant analysis are performed on the same data and compared. CHAID is recommended for consideration by any researcher or planner seeking to optimize the use of scarce resources.

New Software for Market Segmentation Analysis:
A Chi-Square Interaction Detector

INTRODUCTION

Over the last twenty years, a revolution in statistical methods has transformed marketing research. Multivariate techniques first developed in the social and biological sciences are now accepted as indispensible for applied research in both product and service sectors (including higher education). Each year brings another wave of advances that build on old methods and add to the analytical tools that better inform decision makers in a variety of institutional settings.

In two previous papers (Lay, Maguire, and Litten, 1982; and Lay, Maguire, and Nolan, 1982), we explored the practical benefits of a binary segmentation algorithm (THAID). That technique is itself an elaboration of a class of earlier, sum-of-squares based techniques. The first, the Automatic Interaction Detector (AID), was introduced in 1963 by Morgan and Sonquist at the University of Michigan and enjoyed a period of considerable popularity among social science researchers. Unfortunately, some researchers mistakenly viewed AID as a substitute for theoretical modeling. Because of these abuses, AID was criticized and its use declined. Binary segmentation algorithms continue to be used by social scientists, but correctly only as an exploratory technique.[1] Early on, however, one variation or another of this method of segmentation analysis found wider application in business as a marketing research tool, where it

---

[1] See Fielding (1977) for a discussion of how binary segmentation techniques are used by social scientists and pp.85-86 in Sonquist (1973) for a bibliography.

gained much success and general acceptance. In our own research, we found this optimizing algorithm to be a powerful tool for efficiently segmenting applicant and prospective applicant pools. For higher education planning in general, the approach holds considerable promise as a guide for distributing scarce resources to maximum effect.

The technique has appeal in applied settings because the:

1. statistical underpinnings are simple and straight forward,
2. intermediate results convey useful information on how segments are identified,
3. solutions are easily displayed and explained to people who are unfamiliar with higher-level, multivariate models--accompanying statistics are simple means and percentages,
4. results yield information about specific segments that other methods often miss, and
5. policy actions suggested by the analysis can be easily tied back to specific, nonoverlapping segments.

AID AND CHAID

The Advantages of CHAID

See Sonquist (1973) or Lay, Maguire, and Litten (1982) for a discussion of the advantages and disadvantages of AID. CHAID shares all of the advantages of AID over more traditional methods. Both have the facility to find interacting influences other techniques often miss. Both can detect interactions among a much larger number of predictor variables (up to 40) than can other methods. And both can assess the magnitude of effects on specific subgroups making them ideal, as an analytical tool, for planning a differentiated strategy (e.g. a targeted communications strategy).

In addition, CHAID has advantages that make it superior in some important respects to AID (and by logical extension to THAID as well). CHAID uses a chi-square statistic to measure the strength of association between a given predictor and the criterion to be maximized. Because the chi-square statistic is nonparametric, CHAID requires fewer assumptions than AID, which itself makes only minimal measurement assumptions. A number of the advantages of CHAID stem from its use of a chi-square optimizing procedure.

One shortcoming of AID is a selection sensitivity to skewness in the distributions of predictors. To reduce this threat of invalidity to the solution, predictor variables should first be transformed to "even" their distributions before input to AID. Since chi-square statistics are not affected by marginal distributions (as long as expected cell sizes are sufficiently large). [2] CHAID requires no distributional transformations before input.

AID also suffers from a tendency to select predictors that have relatively more categories. To do AID, one should first collapse categories so that all predictors have about the same number of values. This precaution is not necessary before executing CHAID.

CHAID automatically adjusts the chi-square statistic calculated at each selection step with a Bonferroni multiplier, varying in magnitude depending upon the number of degrees of freedom associated with each predictor. Thus, in selection, one predictor is never artifactually favored over another simply because it may have more scale values. Because CHAID requires a categorical criterion variable--a necessity for computing rela-

---

[2] CHAID lessens the problem of small or empty cells by collapsing categories that are sparsely occupied.

tive chi-squares--CHAiD has wider applicability than AID. CHAID can accommodate categorized ratio or interval, ordinal, or even multilevel nominal variables as the criterion to be maximized. CHAID seeks a solution that maximizes prediction of the *distribution* of responses, ordered or not. In practice, then, CHAID can find a segmentation structure that optimizes *any* criterion.

A chi-square test is also used to determine the best division (or "split") on a predictor's scale of values. As a consequence, rather than being artifically limited to binary splits, CHAID may make multiple splits in a single step (iteration). As will be seen in the examples to foilow, this is one of the more attractive features of CHAID.

For readers familiar with log-linear techniques, it should be obvious that CHAID shares many of the same advantages those techniques have over traditional parametric, regression techniques. The similarity between the two nonparametric approaches is more than just superficial. Magidson (1982) has shown that CHAID can be a good first step in a log-linear analysis. One limitation of log-linear modeling is that it becomes cumbersome to analyze more than a small number of variables at a time. From a large set of eligible preditors, CHAID can identify a small subset that may be analyzed in a log-linear equation. Additionally, Magidson shows that CHAID can locate interactions that lead to more parsimonious solutions than one might otherwise find. In turn, log-linear models may be estimated to specify interactions identified by CHAID and test them for statistical significance.

The software to execute a CHAID solution has a number of distinct advantages: First, it is a stand-alone routine that is easily learned and can be called by a more complete statistical analysis system. Second, the

information supplied at each step of the routine is extensive, which makes the procedure transparent and intermediate results quite useful.

Last and not least among these listed advantages is that CHAID is available to any researcher who has access to an IBM computer, either on campus or through a communications network such as Telenet or Edunet. Assistance is available from and the program itself is supported by the developers of the software.[3]

Disadvantages of CHAID

CHAID also shares many of the disadvantages of AID. First, both are exploratory techniques and both are atheoretical. Statistical models built on *a priori* theory are to be preferred, but are rarely elaborated to this level of analysis. In the absense of good theory, these exploratory techniques can still yield important practical benefits and suggest a structure of interacting influences among variables. The techniques search through hundreds and even thousands of cross tabulations much the way a good researcher might, looking for significant patterns and relationships.[4]

Second, both techniques, because they are stepwise procedures, may capitalize on chance variations in sampling. It is desirable to use large samples (over 1000 observations) to reduce the likelihood of idiosyncratic

---

[3] The Software was developed by William D. Perreault, Jr. of the University of North Carolina at Chapel Hill and is available through Statistical Innovations, Inc., 55 Wheeler Street, Cambridge, Massachusetts 02138. A 184 page documentation manual is available to preview the procedure.

[4] To illustrate the complexities involved, consider that with ten predictors, each with three categories, that there are 59,049 possible segments (three to the tenth power).

solutions. Replication is also desirable to confirm the reliability of the branching structure. And, although statistical tests are performed at each step, CHAID does not have an overall test of significance for the final structure. Consequently no single optimal solution is guaranteed. All stepwise techniques share these disadvantages.

CHAID has one disadvantage not shared with AID. The criterion (that is, the dependent or endogenous variable) to be maximized must be categorical. Thus, the tradeoff for the advantages listed in the last subsection is made against the potential loss of information from categorization of ratio or interval level variables. Fortunately, the loss of information is usually small because CHAID can accomodate up to 15 categories in the criterion.

The chi-square statistic, because so few measurement and distributional assumptions are made, is conservative in rejecting a null hypothesis of no relationship. The risk for the researcher, then, may be of too high a chance of Type II error, i.e. failing to reject when a true relationship exists. In applied settings, the cost of a missed opportunity may be quite high. To lower the chance of this type of error, CHAID can be made more or less sensitive by manipulating the parameters that control the selection algorithm.

METHODOLOGY

Data Sources

Data are from responses to a questionnaire administered in a mailing with two follow ups in the summer of 1981 to random samples of 3000 non-applicants (35% response) and to 2000 applicants (75% response) accepted for freshman admission that fall. After comparing sample and known pop-

ulation distributions, we concluded that the difference in responsiveness between samples was the only major bias. Responses were weighted to properly represent the populations from which they were drawn.

The object of segmentation analysis in this instance is to identify prospective students who have some chance of being admitted and who may enrich the entering freshman class. With this purpose in mind, only "qualified inquiries" were included in this study. A qualified inquiry is defined as a prospective student who ranked in the top ten percent of his or her high school graduating class and who scored 1000 points or higher on math and verbal SATs combined.

To yield comparable results, the same small number of variables that can be measured from responses to a "Bulletin Reply Card" were used in the application of each method.[5] These variables were chosen because they could identify segments for particular treatment in a marketing plan to encourage prospective students to apply. Table 1 displays the definitions of these variables.

Insert Table 1 about here.

Methods

The results of three methods will be compared and contrasted on the same data and using the same criterion and the same set of predictors.

---

[5] See Lay, Maguire, and Nolan (1982) for a discussion of the use of a "reply card" to gain information early on inquiries and of why inquiries should be "qualified" by the researcher.

Table 1 Eligible variables for segmentation analysis

Variable Name	Value Labels
Sex	1 = male, 2 = female
Race	1 = White, 2 = Black, 3 = Hispanic, 4 = Asian, 5 = Native American, 6 = other
Geography	1 = Boston area, 2 = rest of Massachusetts, 3 = rest of New England, 4 = rest of Northeast, 5 = rest of United States, and possessions
High School Type	1 = public, 2 = parochial, 3 = private
High School Percentile	1 = 99th, 2 = 95th-98th, 3 = 90th-94th, 4 = 80th-89th, 5 = 50th-79th, 6 = 1st-49th
SATs	1 = 400-799, 2 = 800-899, 3 = 900-999, 4 = 1000-1099, 5 = 1100-1199, 6 = 1200-1299, 7 = 1300-1600
Preferred Program	1 = Arts and Sciences, 2 = Management, 3 = Nursing, 4 = Education, 5 = not offered at Boston College, 6 = undecided

Because the criterion is a nominal level variable, THAID was chosen as the appropriate aid procedure.[6] Then, CHAID will be used to segment the criterion, application yield, in two forms: 1. as a dichotomous variable (1=did not apply, 2=did apply), and 2. as a variable with three values (1=did not apply, 2=applied but did not accept offer of admission, and 3=applied and enrolled). The second form could not be optimized by THAID (one of its limitations) and is presented here to illustrate how CHAID can optimize a process, in this case admissions flow.

Third, a multiple discriminant analysis was performed on both forms of the criterion. Because the second was found to be more complete, only that analysis will be presented here.

RESULTS

THAID Solution

The THAID tree diagram is displayed in Figure 1. This solution is discussed in more detail in Lay, Maguire, and Nolan (1982). The criterion being optimized is application yield among qualified inquiries. The algorithm begins by finding the best binary split of the overall group, estimated here to include 12,120 prospective students (this estimate is the result of weighting sample responses in the correct proportion to be representative of the population). The probability of someone applying after making an inquiry is calculated to be .17, which is simply the relative frequency of applicants to all qualified inquiries.

---

[6] THAID is a procedure in the OSIRIS III statistical package (Institute for Social Research, 1973).

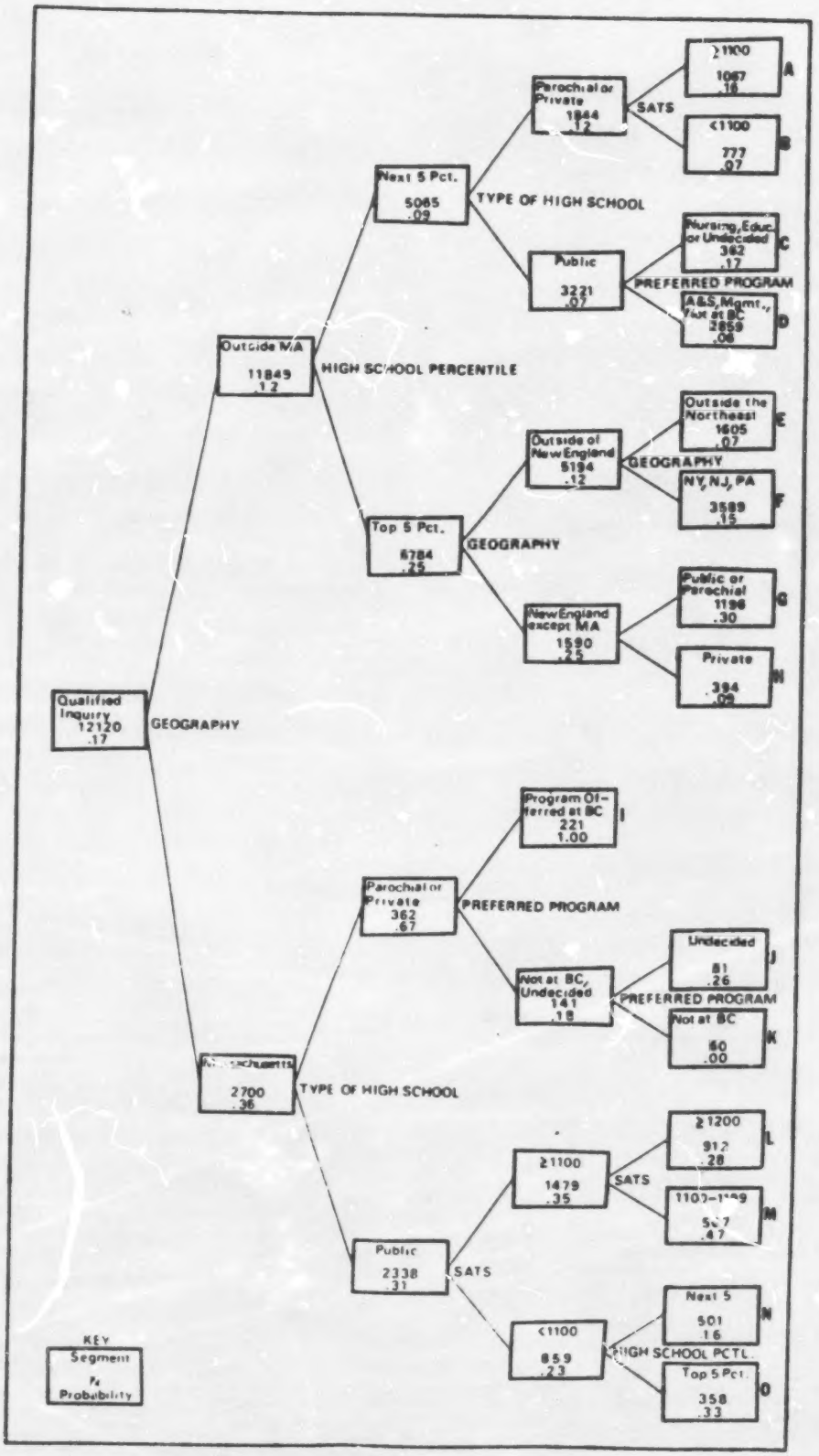


Figure 1. THAID Segmentation of Qualified Inquiries according to Application Yield

Insert Figure 1 about here.

The first variable selected, the characteristic that shows the largest weighted difference on application rates, is "geography." On this characteristic, the best split is between those whose home is outside Massachusetts (determined by combining categories in every possible combination) and those whose home is inside Massachusetts. These results show that the probability of applying is much higher for the 2,700 Massachusetts residents (p=.36) than for the 11,849 out-of-state residents (p=.12).

At the next step of the algorithm, THAID is applied independently to the two segments. In two independent steps, different characterisics are identified as best segmenting each segment. For Massachusetts residents, "type of high school" is chosen, and for out-of-state residents, "high school percentile." That different variables are selected suggests an interaction with "geographic location." Among Massachusetts residents, parochial or private school students are more likely to apply (p=.67) than are public school students (p=.31).

Yet among out-of-state residents, the identified pattern is different. Those who perform well in high school, as measured by relative rank, are more likely to apply than those who perform at slightly lower levels. Of course to determine if a statistically significant interation does exist, the rest of the solution should be examined to see if the branches include the same variables, in a different order (indicating no interaction). Also, interactions may be specified and tested in a log-linear model.

The algorithm is applied in this manner iteratively until preset stopping criteria are met. In this solution, 14 segments, (A through O) are identified, the result of a series of binary segmentations.

The CHAID Solution: One

Figure 2 displays the results of the comparable CHAID solution. Because weighting is done differently in the two procedures, the number of estimated qualified inquiries is higher (n=14.419) in the CHAID solution. The CHAID estimates are more accurate.

Insert Figure 2 about here.

As in the THAID solution, CHAID selected "geography" as the best segmenting variable, based on the largest adjusted chi-square statistic. Even at this first step, the difference between THAID and CHAID becomes apparent. CHAID selected four, distinct geographic segments instead of two (the limit for THAID). In the CHAID solution, out-of-state residents were subdivided to form three, significantly (statistically) different segments.

Among Massachusetts residents, CHAID selected "type of high school" just as THAID did. Again, however, CHAID was able to identify more than two different segments. And for the policy maker, the large difference in application rates between parochial and private high schools in Massachusetts is certainly not trivial, and may be very important to

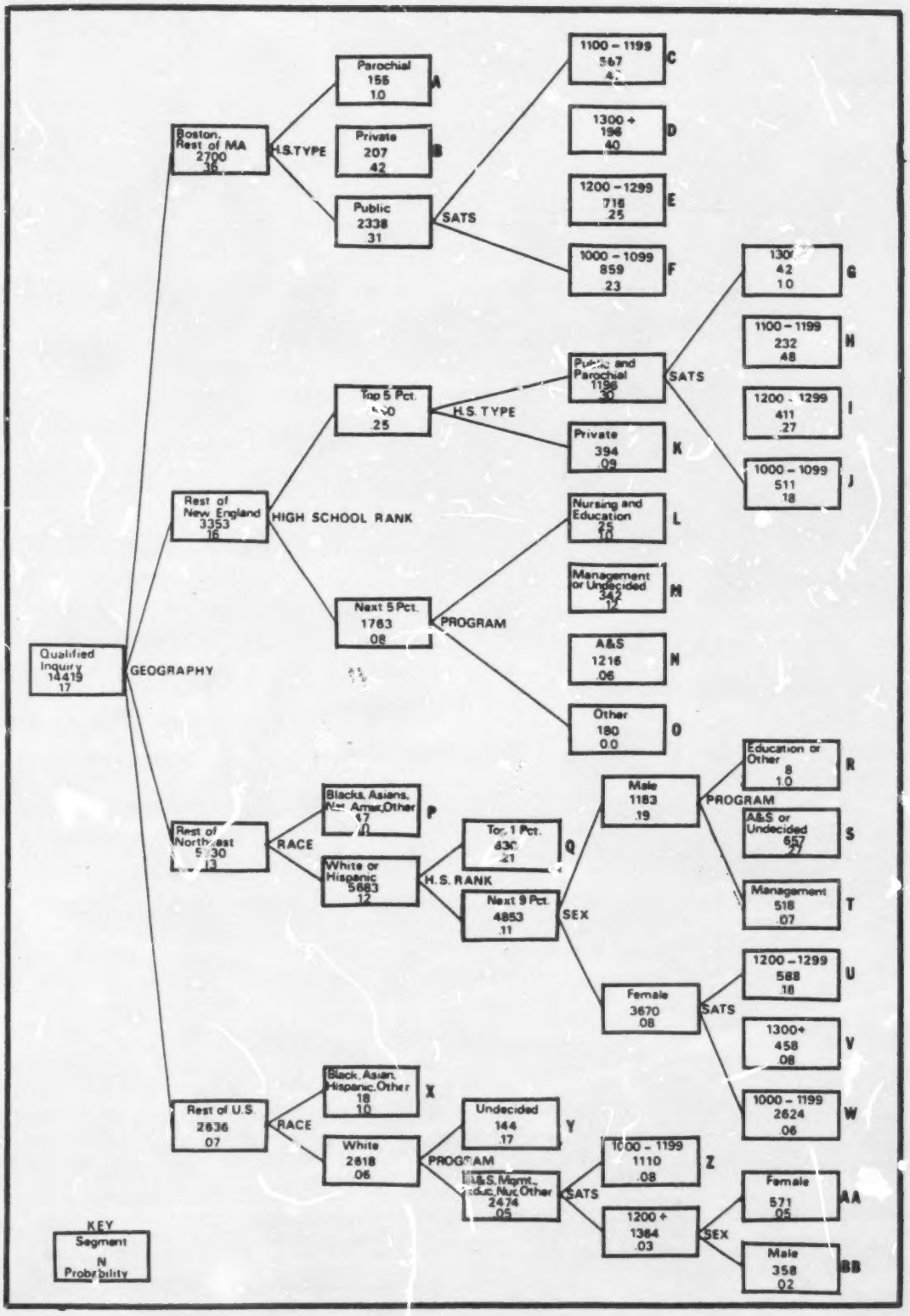


Figure 2. CHAID Segmentation of Qualified Inquiries according to Application Yield

know; qualified inquiries from parochial high schools are much more likely to apply than are those from private high schools.

Because CHAID has the flexibility to choose any number of segments at each step, it is difficult to compare the THAID and CHAID beyond the second step. It seems clear, however, that the two solutions-- the variables being selected and the interactions being identified-- are quite similar. Yet because more splits are made, the structures look different and result in different numbers of final segments.

THAID tends to yield more even-sized segments, which is usually desirable. CHAID is more likely to isolate a small subgroup as a selected segment. This can be a nuisance for the segmentation marketer, since segments that are small in size do not usually make good target groups. The software might be more helpful if it allowed the user to set a minimum size for a segment, but it does not. It can be argued, on the other hand that the researcher should be aware of such "splinter" segments precisely because they are so different from the larger segments. Of course, the marketer may simply decide not to include small segments in a differentiated strategy.

CHAID Solution: Two

CHAID is not limited to a nominal criterion variable with only two values. In the two segmentation solutions just discussed, the variable to be optimized is application yield: either a qualified inquiry applied or he or she did not. The second CHAID solution, presented in Figure 3, was constructed to optimize a *distribution* of possible outcomes rather than a single outcome. The criterion in this analysis is an ordered variable indi-

cating first if a qualified inquiry applied or not and second, for those who applied and were accepted, if he or she chose to enroll. The variable to be optimized, then, takes on three values: 1=nonapplicant, 2=nonmatriculant, and 3=matriculant.

Insert Figure 3 about here.

The shaded areas in the boxes drawn for each segment sum to the equal proportion of qualified inquiries who applied (see Figure 3) so that the two CHAID solutions may be compared. Note that the solutions have only minor differences and that for practical purposes, either could be used to inform a segmentation strategy. Figure 3, however, does give additional information that may be useful about some segments. For example, members of segment L (in Figure 3), including Blacks, Asians, Native Americans, and other minorities, are very likely to apply (100% do), but only a small proportion are likely matriculate (13%).

The Multiple Discriminant Analysis

A traditional discriminant analysis solution is presented in Table 2. As for the second CHAID solution, the criterion is defined by the three qualified inquiry subgroups: applicants, nonmatriculants, and matriculants. Predictor variables were expressed either as interval level variable (SAT scores and high school percentile) or as dichotomous variables. Two functions were identified. The first function distinguishes applicants from nonapplicants. As observed in the THAID and CHAID solutions,

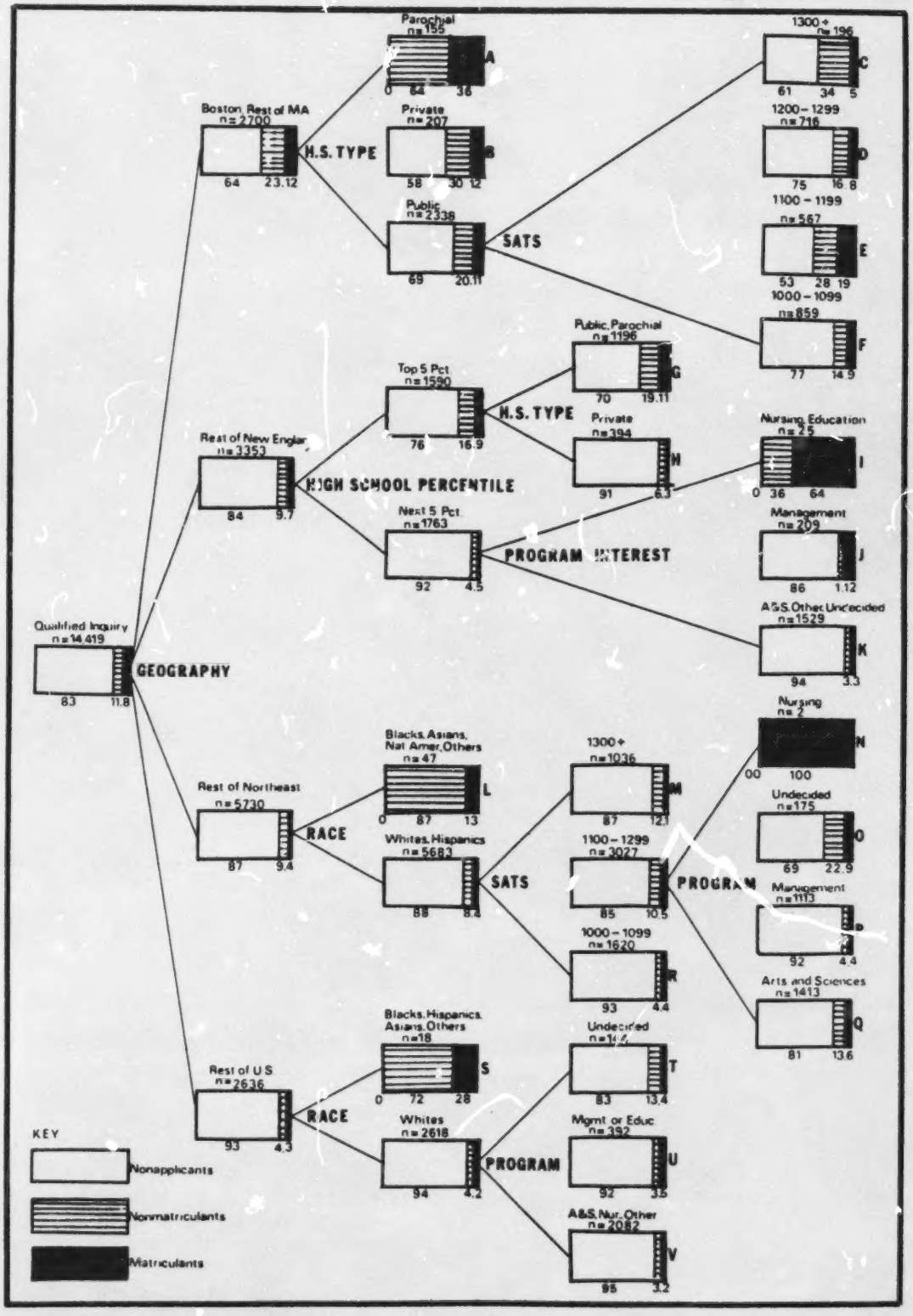


Figure 3: CHAID Segmentation according to Percentage of Qualified Inquiries Who Apply and Percentage of Accepted Applicants Who Matriculate

"geography" (measured here by a Massachusetts/outside of Massachusetts dichotomy) is the best discriminator overall. As did THAID and CHAID, the discriminant analysis identified SAT scores as an important discriminating variable (nonapplicants tend to score highest), although the segmentation techniques detailed specifically how SAT performance interacts with other characteristics of the prospective student.

Insert Table 2 about here.

The second function discriminants between nonmatriculants (those who chose not to enroll) and the other two subgroups, matriculants and nonapplicants. Again, geography is an important factor--Massachusetts residents are more likely to apply (Function One), but less likely to enroll (Function Two). This pattern was seen in the second CHAID solution (Figure 3) in the first segment selected; note that 45% of Massachusetts residents apply, yet only about one third of these enroll.

That multiple discriminant analysis (MDA) has a different objective should be understood here. MDA creates discriminant functions from a set of specified variables. These discriminant functions may then be used to classify a new sample of individuals as likely to be a member of one group or another. For example, Function One could be used to classify prospective students as likely applicants or nonapplicants. MDA in this instance is helpful in predicting overall potential for application among qualified inquiries. It only shows, however, how characteristics combine across the entire sample. MDA does not suggest how subsets of students might be given special attention.

Table 2. Discriminant Analysis of Three Qualified Inquiry Subgroups

Function one (applicants vs. nonapplicants)

Level	Characteristics	D[b]	Group Means
.7	Massachusetts resident	.79	Matriculants .65
-.4	SAT scores	-.45	Nonmatriculants .42
-.5	Interest in program not offered	-.56	
			Nonapplicants -.10

Function Two (nonmatriculants vs. matriculants and nonapplicants)

Level	Characteristics	D[b]	Group Means
.5	Massachusetts resident	.54	
	Minority	.52	Nonmatriculants .64
.4			
.3	Interest in Arts and Sciences	.39	
	SAT scores	.37	
	Interest in program not offered	.35	Matriculants .11
-.2	Female	-.23	Nonapplicants -.09
	Graduated from public high school	-.23	
-.3	High school relative rank	-.31	

[a]Stepwise Rao Discriminant Analysis.
[b]Varimax rotated discriminant function coefficients.

While this examination of a multiple discriminant analysis solution is admittedly brief, it should be enough to illustrate two points. First, MDA provides information that is different from that obtained by segmentation techniques. MDA measures the effects of variables across a population, while THAID and CHAID identify effects for progressively smaller subsamples. Each approach has utility, but THAID and CHAID provide the amount of detail needed to inform a differentiated policy. MDA (and by extension other traditional techniques) are more appropriate for a policy that treats everyone in a uniform manner.

Second, there is enough correspondence between the results of MDA and the two segmentation techniques to say that the same general conclusions might be drawn from either. Given the very different assumptions and objectives of these data analytic techniques, this rough correspondence can only add to the confidence in the findings of both approaches.

IMPLICATIONS

Institutions of higher education face a challenge to better meet the needs and preferences of prospective students and do so within situations that impose severe fiscal constraints. Under those circumstances, each action must be determined to have a reasonable chance of success and resources must be targeted to have the maximum desired effect. Techniques such as THAID and CHAID are methods that better fit the growing requirement that efforts be focused. Traditional techniques that measure overall impact may lead to the selection of policies that are applied wastefully to all and miss lower-cost, targeted policies that yield the same or higher benefits.

CHAID is a technique that is well suited to address these new needs. The statistical basis for this algorithm represents an important improvement over earlier AID methods. CHAID affords additional flexibility, both in the variety of variables that may be optimized and in the way the technique can identify multiple segments at any given step. CHAID is recommended for consideration by any researcher or planner seeking to optimize the use of scarce resources.

Armstrong, J. Scott and James M. Andress (1970), Exploratory Analysis of Marketing Data: Trees vs. Regression, *Journal of Marketing Research*, Vol. 7 (November), 487-492.

Fielding, A. (1977), Binary segmentation: the automatic interaction detector and related techniques for exploring data structure. In C.A. O'Muicheartaigh and C. Payne (eds), *The Analysis of Survey Data*,Vol.1, New York: Wiley, 1977.

Green, Paul E. (1978), An AID/Logit Procedure for Analyzing Large Multiway Contingency Tables, *Journal of Marketing Research*,Vol. 15, pp. 132-136.

Institute for Social Research (1973), OSIRIS III, Release 2. Ann Arbor: University of Michigan, 1973.

Lay, Robert; John Maguire; and Larry Litten,(1982) Identifying Distinctive Groups in a College Applicant Pool. *Research in Higher Education*,95-208.

Lay, Robert; John Maguire; and Charles Nolan (1982), Assessing Market Potential in the Inquiry Pool, paper presented at the Twenty-Second Annual Forum of the Association for Insitutional Research, May 16-19, 1982, Denver, Colorado.

Newman, Joseph W. and Richard Staelin (1971), Multivariate Analysis of Differences in Buyer Decision Time, *Journal of Marketing Research*,Vol. 8 (February), 192-198.

Sonquist, John A.; Elizabeth L. Baker; and James N. Morgan (1973), *Searching for Structure*.revised edition, Ann Arbor, Michigan: Survey Research Center, 1973.

Kass, Gordon V. (1980), An Exploratory Technique for Investigating Large Quantities of Categorical Data, *Applied Statistics*, 29, No.2, pp. 119-127.

Magidson, Jay (1982), Some Common Pitfalls in the Causal Analysis of Categorical Data, *Journal of Marketing Research* 19, (November).

Perreault, William D. Jr and Hiram C. Barksdale (1980), A Model-Free Approach for Analysis of Complex Contingency Data in Survey Research, *Journal of Marketing Research*, 17 (November), 503-515.